DOCUMENT RESUME

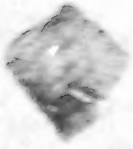
ED 205 533 TM 810 395

AUTHOR Jones, Beau Fly
TITLE Research, Instructional Development, and Practice: The Three Faces of ED.
PUB DATE Apr 81
NOTE 37p.: Paper presented at the Annual Meeting of the American Educational Research Association (65th, Los Angeles, CA, April 13-17, 1981).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Cognitive Processes: *College School Cooperation: Criterion Referenced Tests: Elementary Secondary Education: *Evaluation Methods: *Instructional Improvement: Mastery Learning: Program Evaluation: *Research Problems: *Research Utilization
IDENTIFIERS Chicago Public Schools IL

ABSTRACT

In the first part of the paper, some realistic expectations about schools and curriculum development, the conditions which preclude applying the standards of basic research to evaluate a program of instruction in large school systems, and some new research paradigms which have emerged from practice are defined. In the second part, an effort is made to generate some practical suggestions to modify basic and applied research designs without jeopardizing the integrity of either. In part three, the conditions needed to generate closer school/university cooperation are considered. (Author/BW)



****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED2055533

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

RESEARCH, INSTRUCTIONAL DEVELOPMENT, AND PRACTICE: THE THREE FACES OF ED*

Beau Fly Jones

Chicago Public Schools

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
B. F. Jones
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Ruth B. Love

General Superintendent of Schools

Board of Education

City of Chicago

TM 810 395

*Paper to be presented at the annual meeting of the American Educational Research Association, Los Angeles, April 1981

RESEARCH, INSTRUCTIONAL DEVELOPMENT, AND PRACTICE:

THE THREE FACES OF ED

As school researchers, we heartily applaud the recent movement to direct cognitive researchers in the universities to address school problems so that their research has an impact on schools. However, it is argued that this movement will be limited in its impact, if not aborted altogether, unless there are accompanying changes in research paradigms, the philosophy that defines the goals of research, and the current system of rewards for conducting research.

More specifically, we fear that this new educational researcher (ED) will be schizophrenic because he faces a double bind. On the one hand, he is expected to change student achievement by developing cognitive-based instruction. Yet he is not provided with adequate university instruction or experiences in the schools to define realistically what effective instruction is, how to develop it to apply to school tasks, or how to evaluate it in the schools using the standards of basic research that he has been taught. To make matters worse, the likelihood of obtaining the training and access to classrooms that would facilitate success is limited, given present research philosophies and reward systems. All his life he has been told that the purpose of research is to generate contributions to theory using standards of research generated by the physical sciences. But these research paradigms are not acceptable to schools because they do not address school needs directly. Moreover, they are often inappropriate to assess instruction in schools. Consequently, ED is feared and rejected in the schools. At the same time, if ED tries to be inventive and develop research paradigms that do address school needs, his research may be considered less than adequate in the network of those who make decisions about university promotions, grants, and article acceptances. Thus, ED may actually be punished by his university colleagues for his efforts to achieve a newly-articulated national goal for schools.

What is needed to prevent this schizophrenia is obviously complex and cannot be fully addressed in this paper. However, for several years, the author has been involved with two cognitive-based curriculum research and development projects in the Chicago public schools (CPS). Katims and Jones (1981b) this year are completing the development of the Chicago Mastery Learning Reading materials (CMLR) for grades K-8. These are instructional materials developed for the Chicago Board of Education (1980). Completed materials were delivered citywide in September 1980, and are in use in numerous schools, though we have no way of knowing the exact (or even approximate) extent of their use. CMLR represents a welding of mastery learning theory with concepts from instructional design and cognitive psychology. More specifically, the materials contain information about the structure of text to be read and explicit reading/learning strategy instructions as well as diverse modes of presentation and instruction. Additionally, CMLR is based on a sequencing model which controls for the fact that many students in a given classroom may lack the prerequisite skills for the objective being taught. Having developed these programs from scratch with little graduate training that was directly applicable, we have had to develop some research paradigms and a research philosophy that may be useful to ED.

In the first part of the paper we try to define some realistic expectations about schools and curriculum development, the conditions which preclude applying the standards of basic research to evaluate a program of instruction in large school systems, and some new research paradigms which have emerged from practice. In the second part, an effort is made to generate some practical suggestions to modify basic and applied research designs without jeopardizing the integrity of either. In part three, the conditions needed to generate a school/university marriage are considered; besides making ED "a legitimate child," such an alliance would have many benefits for both partners.

Cognitive-Based Instruction in the Schools: Operational Problems and Solutions

The Politics of Negotiation for Instructional Change

Before ED can change instruction in the schools, he must be admitted. Two true stories illustrate the need for him to gain a realistic picture of school politics and school tasks before he enters into negotiations for school/university alliances. The first story is about a group of ED researchers who came to a nearby school system to offer their services to that system. They had heard that the curriculum there was in dire need of evaluation and repair. With much good will, they offered to help the school analyze the problems in the existing programs, develop revisions and supplementary materials, and conduct research which would identify poor teaching practices. The meeting was very short and ended with the school declining to accept such an alliance in spite of the "free" assistance and products available in the offer.

What the researchers had failed to understand was that their behavior conformed to the typical stereotypes that school administrators often have about university researchers. First, they do not want to have anything to do with the "dirty work" of curriculum development created within the school (as opposed to the lab) but are most happy to come in and criticize after the fact. Second, it was evident that these researchers had no idea of the phenomenal cost of changing even a part of a system that has already been implemented. Third, they did not understand that most of what they wanted to do was threatening to the jobs or status of the initial developers and to school administrators who would have to contend with union pressures against any form of teacher evaluation.

The same group of ED researchers came to the same school system several years later with one simple question: "How can we help you?" Their offer to negotiate was accepted immediately and a committee of school/university

researchers drew up a proposal that satisfied the needs and politics of both partners in the collaboration. The point here is twofold: Clearly, such alliances are most likely to be formed when the school system is invited and encouraged to articulate its needs and the conditions under which technical assistance is useful and doable. Equally important, this group learned by trial and error because the gap in school/university communication and interaction is so great. Ironically, school staff are almost never employed as consultants in universities to provide information about school administration and the politics of changing instruction or about the needs of classroom instruction and the typical teaching/learning environment under which instruction takes place. Perhaps it is time for this type of consulting as well as for instruction within universities from ED researchers who have had on-site experiences.

After gaining admittance to a school system, the novice ED researcher must face a host of problems. At this point, we will shift gears and use our experiences in the Chicago public schools as a means of communicating some of the discrepancies between university research on instruction and learning in the laboratory and the conditions of curriculum development and research for students in the classroom.

Teaching/Learning Conditions

Recently, Brown, Campione, and Day (1980) developed a model for training students to learn from texts. The four focal points in their "tetrahedral model" were (1) the characteristics of the learner, (2) the critical tasks to be performed, (3) the nature of the instructional materials, and (4) the learning activities needed to perform the critical tasks.

While this model is clearly a contribution to research and practice (given the lack of other models), it omits a most important set of factors: the teacher and the teaching/learning situation. Instruction of any kind requires an

effective delivery system which meets the demands of the teaching/learning situation. It is useless to develop an instructional program that only the experimenter or the developer can implement successfully. Consequently, it is crucial to consider the organization of the school system as a whole as well as typical classroom conditions. Moreover, in schools this is the first consideration to be made because the resulting decisions define the parameters of the instructional system.

Three considerations were outstanding in the development of CMLR. First, it was evident from the outset that we had been hired to provide instruction for an existing system of objectives and tests: Continuous Progress/Mastery Learning (CP/ML). This system consists of 1,400 objectives which span the nine years of elementary school that are equivalent to kindergarten through eighth grade. Of these, 273 are defined as key objectives and are assessed by criterion-referenced tests. The existence of these objectives and tests meant that we were not free to devise our own objectives (though in fact we did occasionally add to the existing ones). Their existence also meant that any instructional materials that were developed would necessarily embody mastery learning theory.

Second, there was no way to control the process of implementation. Chicago has about 450,000 students housed in over 500 buildings with that many principals and school staffs. Currently, there are over 23,000 teachers, some of whom are not fully certificated. Although there is an extensive staff development program, compared to most large inner city school systems, it was not designed to provide intensive or extended training for mastery learning. Moreover, while there are several innovative mechanisms for monitoring pupil progress, there are no central mechanisms for monitoring the implementation of a given program of instruction or the quality of instruction in the classroom. Additionally, some schools have freed staff to serve as resource persons for teachers, but it is the bureau heads who

define their responsibilties and there was no opportunity for us to utilize such persons. In essence, although we do train anyone who requests inservice there are no mandated or contractual provisions for site managers or resource persons whom we could systematically train as is the case in DISTAR (Englemann, 1978) and the materials developed by Reid (1978). Additionally, there was a considerable amount of teacher mobility within each school year due to absenteeism, transfer, and temporary assignments to integrated school sites for special programs.

This inability to control what happens in the classroom led us to favor curriculum development, rather than staff development, as the most useful mode of delivery. By developing instructional materials for teachers and students, we could provide the essentials of high-quality instruction directly to the students and reduce teacher variance to a minimum, as is done in good laboratory research. Providing instructional materials made it possible for both substitutes and transfers to maintain continuous instruction with little or no training. Finally, instructional materials can provide a stability of instruction from generation to generation (not to mention from city to city) that no staff development program can provide. For all these reasons, it was concluded that developing instructional materials would be a markedly more effective mode of delivery, compared to staff development, to address the teaching/learning conditions of a large urban system. This conclusion is supported by subsequent studies such as the Rand (1978) report and Joyce and Showers (1979) review of over 200 staff development programs; both studies demonstrated that such programs had little impact on the quality of instruction in the classroom or on student achievement.

A third major problem that was inherent in the organization of the Chicago public schools was the marked heterogeneity and mobility of the students in each classroom. Most classrooms are organized by age cycles and reading levels since there are no grade levels in a continuous progress system. This means that in

any classroom, there may be one to two age groups or more and one to three reading levels. And most schools combine CP/ML with instruction from one of any number of basal programs. Other classroom problems are caused by student absenteeism, transfer, regular classroom reorganization, and suspension. These conditions guaranteed that, for any given objective, there might be a substantial number of students who lacked the prerequisite skills for that objective.

The solution to this problem was the development of a sequencing model on which each unit of instruction "begins at zero." That is, it is assumed in each unit that every child lacks knowledge of the concepts, skills, and vocabulary needed for each objective. Accordingly, the instruction in each unit begins with a heavy teacher input and content that is simple/concrete/familiar/explicit/short, and progresses to independent practice using content that is more complex/abstract/unfamiliar/implicit/long (see Jones, Amiran, & Katims, 1980; Stoll, 1980). Sequencing instruction in tiny increments builds in success for those who lack the requisite skills and minimizes the problems caused by absenteeism and transfer.

Other critical decisions emerged from the nature of the objectives and tests and the characteristics of the learner--including the learner's repertoire of available learning strategies.

Target Tasks and Student Populations

The most immediate problem in considering how to develop instruction for school tasks was the lack of research on instruction that was directly relevant. Apart from Steinberg's (1976) research on analogies, there was little research on instruction for many of the tasks that were identified as key objectives, such as forming contractions and other rules of grammar, differentiating fact from opinion, and separating supporting and nonsupporting facts in an argument or theory. In other instances, there were relevant theories, but there was a

great "distance" between the laboratory task and the school task in terms of complexity, length, and content. Fortunately, there is now a growing body of research on a number of school tasks such as summarizing (Brown, Campione, & Day, 1981), making inferences (Hansen & Pearson, 1980), and finding the main idea (Andre & Anderson, 1978-79). However, in developing any comprehensive curriculum, there are inevitably numerous objectives for which no guidelines are available.

These problems raise the question: How do you develop cognitive-based instruction when there are no guidelines from research or when there are guidelines which are not appropriate for the target population? Rohwer's model was avant-garde. As early as 1971, he anticipated these questions and provided a model for ED researchers (Rohwer, 1971). In a seminal article discussing the importance of developing proficiency skills for low achieving students, he outlined a multi-stage research process for identifying effective learning strategies from high achieving students, verifying the strategies, developing a training program based on the research, and revising that program to meet increasingly more stringent criteria.

These were essentially the stages that were followed in the development of CMLR. However, since we had no capability to conduct these steps using basic research paradigms as suggested by Rohwer, we substantially modified his procedures to adapt them to school conditions. Instead of identifying skilled subjects of school ages to test and interview, we used our own writing staff to do the tasks; we asked them to "think backwards" to identify not only the subskills in the task (a traditional task analysis procedure) but also the learning strategy or thinking process they used to master the objective. The strategies derived from this process were then regarded as a list of hypotheses to be used as guidelines for the development of the instructional materials. These hypotheses were then tested by stringent field test procedures (see below, p. 16).

Another problem not covered by Rohwer's model was the problem of defining what constitutes effective comprehension instruction for low achieving students. Initially, we began with the notion that providing explicit learning strategy instructions (e.g., to visualize, to find the main idea, to infer) constituted effective comprehension instruction as this is defined by Durkin (1978-79). Although we were unaware at the time of other models being developed simultaneously, the model we first adopted (see Jones, Monsaas, & Katims, 1979) was similar to those developed by Brown, Campione, and Day (1980) for learning from text; as well as models of direct instruction (Carnine, 1978; Rosenshine, 1979) which involve both guided and independent practice. This model was quite effective for some study skill strands (e.g., structural analysis of words, learning the different parts of a book, and so on). But it was quite inadequate for high order study skills such as summarizing and outlining, given the characteristics of the learner. Early field testing indicated that direct instruction by itself did not fully address the needs of the low achieving students; to them the learning strategy instructions were <u>additional</u> behavioral objectives; i.e., it was not possible to instruct low achieving students "to find the main idea sentence" without additional instruction in the process of how to do it. In the absence of research on this problem, we went to the field.

Extensive analysis of piloted materials, interviews with students and teachers, observation of good teachers, and re-analysis of the literature led to an important discovery: the most successful instruction in the classroom identified the <u>structural characteristics</u> of the text to be read. Main idea sentences, for example, vary in content and location according to the type of paragraph (e.g., sequential vs. compare-and-contrast). Good teachers provided this information to the students and showed them how to apply it to the new paragraphs by "thinking aloud." We concluded that in order to teach a given learning strategy, we would have to provide structural information about the

text to be comprehended; i.e., information about the structure of the text; text being defined quite loosely to include word problems, grammar rules, and the like; as well as prose texts such as compare-and-contrast and sequential texts. Accordingly, to the learning strategy component (LS) of our model we added an additional component, structural information (ST), plus step-by-step instructions in the process of applying LS and ST to diverse text examples. Self-addressed adjunct questions were inserted to prompt students to use the strategies and structural information (see Jones, Amiran, & Katims, 1980).

The next major problem was evaluation.

Program Evaluation

It was evident from the outset that the nature of the instructional program was far too complex to obtain anything other than gross measures of its success as a whole; clearly, we would have to leave the assessment of component parts to others. It was also apparent that the teaching/learning conditions would present numerous problems of measurement. Especially problematic were (1) the inability to assess CMLR materials apart from the implementation of other parts of the CP/ML system which is constantly evolving, (2) the inability to assign teachers or students to groups, (3) the lack of control over the quality and length of instructional time, (4) the inability to control (or even assess) the multitude of CMLR/basal combinations times level of students times age of students times teacher variations without daily monitoring of each classroom, (5) the inability to control teacher variations due to absenteeism and differences in the style and use of the materials in the classroom, and (6) the fact that mastery learning procedures eventually yield highly skewed curves with increasingly low variances, a distribution that is not really appropriate for analysis with traditional statistics (see Note 1).

Moreover, even with daily monitoring in the classroom, the variations may be tremendous. Consider the six classrooms used to field test the first set

of upper age cycle materials (Chicago Mastery Learning Reading materials with Learning Strategies, CMLR/LS) in 1977-78. Teaching style varied from essentially self-directed instruction to teacher-directed instruction. Additionally, there were problems in teacher transfer and absenteeism. In one field test (Jones, Monsaas, & Katims, 1979) using six classrooms, the following events occurred. Teacher 1 missed school for several weeks due to a car accident and eventually transferred. Teacher 2 was ill and/or absent for most of the year; her eighth graders, anxious to graduate, virtually taught themselves. Teacher 3 refused to follow the prescribed mastery learning group instruction and used instead the traditional ability groups with basal instruction and persisted in giving all students the correctives because she thought they were effective. She transferred two months before the annual Iowa Test of Basic Skills (ITBS) leaving the students with a series of ten substitute teachers, all of whom seemed a poor match for the students. The eleventh appeared to be an effective teacher by all counts, but he was removed to uphold the desegregation mandate. Teacher 4 was new and very inexperienced. Teachers 5 and 6 were average but differed markedly in style. Needless to say, none of the class sizes was equal or even constant.

Significantly, these classrooms were not atypical in their variances. Teachers in the control school were absent less often but one of them was clearly a poor teacher while another spent two months preparing her students to take the ITBS and a locally developed test. Another was reassigned to a special program. And so on.

Not surprisingly, the standard deviations for the ITBS scores were extremely high when the classroom was used as the unit of measurement. Consequently, in order to obtain any sense of what the students as a group achieved, we used the students as the unit of measurement in analyses of co-variance (using matched classrooms) even though this was theoretically

inappropriate. This seemed more desirable than dealing with figures which were quite distorted by the high variances (Jones, Monsaas, & Katims, 1978). As a result, some of our behaviorist friends wagged fingers with statements about the inappropriate use of traditional statistics and possible, if not likely, halo effects.

This experience caused us to think about the possibility that traditional statistics were indeed inappropriate for program evaluation in any large school system, regardless of possible halo effects. Considerations of halo effects, the correct unit of measurement, and regression effects simply should not apply to situations in which there will invariably be marked variations in teacher use of the materials; changes in student and teacher populations; student identification problems causing numerous cases to be labeled as "missing data" (typically about one-fourth to one-third of all data yielded by the computer); and confounding of the instructional program with numerous other variables.

Nor was the situation any different after the piloting of the materials. The following summer, the summer of 1978, Chicago Mastery Learning Reading materials with Learning Strategies (CMLR/LS) were used in a special mandatory summer school in which considerable irregularities of instruction occurred as the result of changes in the expected number of students (from 15,000 to 22,000) and in their assignment to summer school sites (from nine to 129). The latter piece of information was learned only just a few weeks prior to the start of summer school and the former was learned one week before the program began. In one of the worst problems of distribution of instructional materials, 8,000 student booklets were lost for several days (see Jones & Katims, 1980). Incredibly, the results of this summer school were most encouraging in spite of these conditions. Of course, the results were clearly an interaction of the new promotion policy, the materials, implementation factors, and halo effects resulting

in the belief that this summer school was so special. The following year involved even more disruptions due to the school's financial crisis and desegregation compliance; literally thousands of teachers were transferred (sometimes to teach age groups and subjects they never had taught before) and many projects were terminated altogether. Also, comparative data for summer school that year were unobtainable.

In spite of these problems, it was evident that these mandatory summer schools were noteworthy in many respects (Jones & Katims, 1980). Yet descriptions of them were unlikely to be communicated outside the oral presentation of a professional paper because of the numerous methodological problems which prevented making inferences about causation. It is also unlikely that these materials would have ever been completed if school personnel used the standards of the behavioral sciences as the criteria to make decisions.

This, of course, is a serious source of contention among some of our university colleagues. They argue that only experimental data or other forms of "clean" data should be reported publicly and that only when such data are obtained should decisions be made about the program's effectiveness and its use in the schools. While we agree that such standards are appropriate for the lab, they cannot be applied to schools. School administrators frequently do not trust or understand statistics. Moreover, they have critical implementation criteria such as the teacher's attitude toward the program and its cost effectiveness. Additional factors include the general tendency to apply the rule of "serving a clientele." According to this rule, if a program is perceived as good, it should be distributed citywide as soon as possible; it is morally (and politically) wrong to restrict a good program to a few.

These problems are hardly unique to the Chicago public schools. To the contrary, Chicago is in the forefront of other outcome-based schools in the development of objectives, tests, and instructional materials as well as in the development and use of monitoring systems and staff development programs.

Many large cities do not even have a Department of Research and Evaluation or any time off for teachers to have inservice. And no school can tell the parents that their child has been selected for instruction in a control group. Moreover, problems such as absenteeism, class size, and management/union disputes causing disruptions of instruction are quite limited in Chicago, compared to other cities. And Chicago is relatively advanced for student achievement in inner city school systems (Benjamin, 1981).

Thus, it would seem that the problems described above are an intrinsic fact of life of inner city schools. If this is true, and if our analysis is correct that traditional statistics simply do not apply under these conditions--especially to the number of schools developing mastery learning systems--what can be done?

One solution was to try to turn the problem into an asset. To do this, we extended the widely used concept from basic research: learning to criterion. It was reasoned that this concept should be extended to apply not only to the students but also to the materials. What this means is that in addition to designing cycles of testing, feedback, correction, and retesting for students, as is the case in learning to criterion and mastery learning, we must also design cycles of testing, feedback, revisions/reconceptualization, and retesting for the materials. This concept of criterion-referenced field testing therefore shifts the burden of responsibility for success from the teachers and students to the curriculum developers; if the students do not pass the tests for which the materials were written, the materials are revised and field tested until the specified criteria for success are met.

Thus, the concept of CRFT stands the principles of mastery learning and learning to criterion on their heads; in addition to using these principles to assess the students' mastery of specific objectives, we have extended them to assess the skill of the curriculum developer in facilitating student achievement under typical teaching/learning conditions. The rationale for this extension is

based on two assumptions: (1) All, or almost all, students can learn to specific criteria if they are given appropriate instruction. Therefore, errors and failures generally must be taken as information about the effectiveness of the materials, and the materials should be changed on the basis of analysis of the errors and failures. (2) The materials must be sufficiently robust to be used effectively in diverse teaching/learning situations. This is because often it is not possible to control the way that the materials are used in the classroom. Accordingly, CMLR materials are field tested under two conditions. First, the participating classrooms must reflect the range of typical teaching/learning situations. In practice, we have found that five or six classrooms are adequate if they are selected to reflect the range of typical variations. Second, there should be a minimum of effort to control how the teacher uses the materials in the classroom, other than providing the materials and a minimum inservice.

Moreover, the same principles that are used to define mastery learning procedures for students may be used to define mastery-based field testing. Specifically, it is possible to set criteria for the success of a given unit in terms of the numbers of students who pass the tests. If a minimum of 60 percent of the students do not pass the formative (diagnostic) test, the curriculum developers analyze the test and student activities preceding the test, using the data from daily classroom observations, interviews with students and teachers, and analysis of errors. This feedback is then used to correct (revise) the materials which are then returned to the field for retesting. The same guidelines are used to assess and revise the remedial activity and second test. If, after retesting the unit, there are less than 60 percent of the students passing the formative test and 80 to 90 percent passing the second test, the materials would be revised and field tested until these criteria were met.

Beyond setting measures of expected success for the field testing of individual units, we are also interested in setting criteria for successful use of the materials in schools over time; i.e., in setting criterion-referenced program evaluation measures. The minimum goal set for reading achievement in the Chicago public schools is a year of reading achievement for a year of instruction. We accept this goal as our minimum criterion for a year's use of CMLR. Thus, for one year of using CMLR materials, we would expect an average of one year's gains on the ITBS (or whatever norm-referenced test is used) for students in each age cycle. In terms of administrative units, we would ultimately expect 80 to 100 percent of the classrooms within a school, schools within a district, and districts within a system to obtain an average of one year's growth on the ITBS during one school year. Until those goals are reached, we can use changes in the percentages as program evaluation measures.

There are also expectations in terms of variance. Mastery learning theory predicts that the correlation between pretest and posttest should decrease as a function of the number of objectives mastered: the greater the number of objectives mastered, the lower the correlation. This is because the feedback and correction procedures in mastery learning equalize the opportunities for success on each succeeding objective so that instead of obtaining a normal curve of distribution for each CRT, the performance curve is increasingly skewed to reflect the high numbers of students who regularly reach mastery. This means that pre- and posttest correlations and various variance statistics (see Katims & Jones, 1981a) should decrease over time in classrooms, schools, and districts in a mastery learning program.

Finally, there are implementation criteria. If the teachers and administrators do not like the materials or find them difficult to implement for any reason and therefore do not use them, they may as well not have been

developed. For this reason, the percent of teachers, schools, and districts using the materials is an important program evaluation measure. Moreover, it is often the case that newly-developed materials survive only as long as the original implementing staff continues to provide support. We wanted to develop instructional materials that were capable of survival and growth in the host school system after the developers have departed. And if other school systems are to use the materials, it is essential to design them so that new implementation groups can take on the functions of the original implementation group. Thus, the final criterion of the program's effectiveness is the successful and continued implementation of the materials by new implementation groups, "successful" being defined in terms of facilitating student achievement.

Thus far, in this section, we have described school criterion-referenced measures that were developed because traditional methods of prediction and assessment were inappropriate for the messy data typically yielded in schools. Additionally, in developing and implementing the CMLR materials, we believe that we have observed important classroom and building effects in terms of the attitudes and behaviors of students, teachers, and administrators as well as in school and classroom climate. We have attempted to discover evidence of these effects in the readily available data from the citywide standardized testing program. In particular, we have tried to look at some traditional statistics in rather untraditional ways. In doing so, we have identified some potential indices of effective instruction and effective building management. There are, perhaps, some methodological problems with these indicators. It should be understood that our purpose is not to propose these statistics as proof of program effectiveness. Rather, we are advancing these indicators for their potential usefulness in assessing classrooms and buildings. We are suggesting that with theoretical study of the properties of these indicators as well as continued practical study, these statistics might turn out to work in the way

they appear to work. And if so, creative statistics, to solve a practical problem, will have contributed to methodological theory.

Before describing these measures, it is necessary to describe briefly the implementation of CMLR at May School, our flagship school. From 1976-79 we collected three sets of data for formative evaluation materials at specific grade levels at other schools (Katims, 1979; Katims, Smith, Steele, & Wick, 1978; Jones, Monsaas, & Katims, 1979). Since we could not conduct any experimental research for reasons stated above, we decided to focus on studying the implementation of CMLR at one school from kindergarten through eighth grade: the May School, located in an all-black area characterized by low income and high crime rate. In the year before our implementation, the school was organized in terms of the reading levels for one of the basals. It was in the fifth percentile in terms of citywide reading scores. And it had numerous problems of low teacher morale and absenteeism as well as student discipline, having about 200 suspensions per year for about 1,600 students. In 1978, a new principal and a building reading coordinator were assigned to May. In the fall of that year, they reorganized the school to implement CP/ML reading levels, objectives, and tests, using the available basal readers as the instructional program in most classrooms. (All basals used in Chicago are correlated to CP/ML objectives.) Additionally, they decided to try a bottom-up implementation of CMLR materials using four teachers. If these four teachers were successful, there would be a full-scale implementation the following year. The reading gains in the four classrooms were 1.2, 1.4, 1.6, and 1.1 years. Thus, in 1979-80, all teachers were given CMLR materials.

At the end of the 1980 school year, the principal at May School calculated the mean gain in reading achievement for each classroom and rank-ordered the classrooms from highest to lowest. As he did so, he noted that 15 of the 32 classrooms had gains of one year or more. He was very pleased with this data.

Another eight rooms had gains between seven months and a year, and he was satisfied if not overjoyed with this progress. But the remaining nine teachers were placed on a list for further work because, according to him, they had made little effort to use the mastery learning program effectively.

Yet, as researchers we knew that without systematic classroom observations, we really couldn't say that these were non-mastery learning classrooms. These might well have been the worst teachers, and it is possible that the CMLR materials simply did not help the worst teachers. The problem was that we did not have any control data to separate teacher effects from mastery learning effects.

So we had to invent control data. We realized that both the top group of teachers and the bottom group of teachers he identified had not used mastery learning materials or techniques to teach mathematics. And since these elementary classrooms are self-contained, with each teacher teaching both arithmetic and reading, we could use each teacher's math instruction as a control for mastery learning reading instruction.

Therefore, classroom mean gains for both reading and math were calculated for eight of the top ten and for the bottom ten classrooms. (Some classrooms are not included because pupils were too young to have a test score for the previous year, the prescore.) When these means are compared (see Table 1), the results are very revealing. Whereas the reading gains show the top teachers to be nearly four times as effective, the math gains portray two nearly equivalent groups of teachers. Thus, this arrangement of test data shows that mastery learning teachers are more effective than teachers not using these materials or methods. But more importantly, the data reveal that teachers using mastery learning techniques are nearly twice as effective as they are using traditional instructional techniques.

- - - - - - - - - -

Place Table 1 about here.

- - - - - - - - - -

Of course, there are a lot of problems with using test scores and gain scores as measures of teacher effectiveness. In addition, it is apparent, from attempts to derive these same comparisons in non-mastery learning schools, that these two classroom statistics can reveal radically different curriculum emphases: one school included a classroom with virtually no reading instruction and one classroom with no mathematics instruction. Still, the results of this analysis confirmed the principal's hunch that these teachers had not participated in the program effectively. As school educational researchers, we intend to keep track of this statistical comparison.

In looking at the results from the building as a whole, it is even more difficult to identify clearcut evidence of the effect of mastery learning. Partly, this is due to the fact that all of the teachers and all of the students were not fully involved in mastery learning instruction. Thus the effects are only those of a partial implementation. The biggest problem, however, is defining adequate background data for comparison purposes. Other schools are not quite comparable to May, and May's prior history is itself confounded by yearly changes in the school and the school system.

Fortunately, the theory of mastery learning, because it makes clear predictions about the nature of outcomes, provides some clues as to what to look for. Particularly, the theory specifies treatments designed to minimize the effects of prior learnings and abilities. Thus, mastery learning instruction is supposed to generate (1) higher mean achievement, (2) lower variance of achievement, and (3) reduced, if not eliminated, correlation of gain to prescore. The effects of mastery learning instruction are more equally distributed over classrooms, and we generalized that the effects of mastery learning schooling would be evidenced in the some sorts of changes in building statistics, across classrooms and students.

First, we looked at the distribution of classroom gains and at related classroom statistics. We reasoned that the effective use of CMLR in at least three-quarters of the classrooms should have affected this distribution in the same way that mastery learning instruction leading to 80 percent pupil mastery affects within-classroom statistics.

These results are precisely what we discovered. The proportion of classrooms making gains of one year or more increased from 37 percent in 1979 to 47 percent in 1980; similarly, the mean classroom gain increased from .74 to .90 years. This increase in gains was accompanied by a decrease in the variance of gains from .28 years to .21. And the correlation between classroom gains and classroom prescore dropped from a statistically significant .35 to a nonexistent .03. Before we can say much more about these statistics and changes in statistics, we will need to know a great deal more about the general properties of them. However, it was gratifying to discover that these gross indicators of instructional success do behave in the manner and direction which we had predicted on the basis of the theory.

One of the aspects of the May School experience that impressed us was the effective use of the instructional management system by the principal and building reading coordinator. These two met with teachers on a biweekly basis to review the classroom progress and to discuss/solve individual problems. We have argued that this type of activity is essential in an effective school and we were interested in identifying the effects of such management in the school statistics. Primarily, we expected effective building management to be evidenced again by both an increase in the overall effectiveness of the school as well as in greater consistency of this increase (reduced variation).

The statistics used to describe schools in Chicago are the mean ITBS grade equivalents for the various ages of students in the school: age 7 through 13.

Thus, for each school for each year there are seven age-means in reading ability. Gains of each age group can therefore be obtained by comparing two successive years of scores (e.g., 1979, age 8 = 7.8; 1980, age 9 = 8.6; the nine-year-olds gained eight months in 1980). Looking at a school as a whole in terms of the six age groups for whom there are scores from both years is therefore a good indicator of the school cutting across classroom effects.

We used these six gain scores for the school for the year as an indicator of the school effect. We obtained the mean of these means as a measure of gain, and we used the standard deviation of these six numbers as a measure of variation of effect (see Table 2). In doing so, we observed the same type of mastery learning effect. The gain, as determined by this method, went up, while the variance went down. This result led us to believe that we could see in these gross statistics some analogies to traditional, experimental indicators.

- - - - - - - - -

Place Table 2 about here.

- - - - - - - - -

To try to get additional perspective on the meaning of these indices of building effectiveness and consistency, we calculated similar mean gains and variances of gains for approximately 100 Chicago schools. We arranged these in a scatter plot with mean gain and standard deviation of gain being the axes. What we saw was that there was a citywide relationship between gain and variance. It is easy to produce a small variance of effect; most schools which produce small effects do so uniformly. But as the mean gain increases, so does variance, indicating that great gains were being made by some groups of students, but not throughout the building.

In fact, when we added the May School 1979 data into the scatterplot, we saw that the school was near the center along a diagonal of rising means and variance. But the 1980 statistics put May sharply off the diagonal. This

conforms to our intuitive understanding of the relationship between these two statistics. It is difficult to know if a school with a mean gain of .91 and a standard deviation of .29 is better or worse than a school with a gain of .98 and a variance of .33. But a school with a gain of .91 and a variance of .58 seems much less managed and less effective for all of its students. Moreover, these decreases in variance are particularly significant in the light of the typical classroom conditions described earlier. If this "quick and dirty" statistic can be determined to reveal the school practices which we believe that it does, this may be a practical tool for district administrators' needs as well as a useful statistic for research.

To summarize, cognitive-based instruction in the schools has numerous operational problems for prospective ED researchers: recognizing that all employees of the school system work to attain its goals, not the goals of theory; addressing the problems in the teaching/learning situation which includes heterogeneous classrooms and lack of control over the implementation process; devising effective research and development paradigms to deal with unresearched school tasks as well as the special needs of low achieving students; and finally, generating useful measures of assessment for the field test process and the effects of the materials in schools.

Various solutions were suggested. Of particular importance was accepting the responsibility for modifying and testing the instructional materials until they prove to be effective by pre-established criteria. The concepts of criterion-referenced field testing and criterion-referenced program evaluation were developed in response to these responsibilities. Also, given that traditional statistics are inappropriate to apply to school data because of their variances in variance (i.e., high variances with no program and low variance with good program implementation), it is important to consider new ways to use traditional statistics and/or to develop some innovative statistics.

Specifically, a reading/math paradigm and two variance statistics were examined.

Methodological Issues

In the previous section, all of the problems and solutions were operational. In this section we would like to consider some philosophical issues that underlie the paradigms that are traditionally used to develop cognitive-based instruction. Throughout this section these issues will be examined from the perspective of practice. The aim here is not to integrate the three faces of ED (research, instructional development, and practice), as our analogy to the three faces of Eve might suggest. Rather, our goal is to suggest ways in which both basic and applied research might address the problem of developing and evaluating cognitive instruction so that research is more relevant to practice.

Basic Research. This type of research begins with a hypothesis derived from theory and uses scientific methods to conduct well-controlled studies in the laboratory. The "teacher" in such experiments is usually the experimenter, who implements the instruction exactly as specified. None of these characteristic features of basic research can be changed without altering the nature and therefore value of this research. At the same time, there are several decisions which could be modified so that basic research is more useful to schools: (1) extend existing research on school relevant tasks (e.g., the variety of tasks requiring inferences), (2) select school relevant tasks for new areas such as critical thinking (e.g., discriminating facts and opinions), (3) where possible, use less contrived tasks, (4) develop instructional treatment conditions that are effective for low achieving students, and (5) document the effects of some school factors that are dysfunctional to learning (e.g., the effects of poorly written texts). Clearly, these suggestions reflect guidelines that are already adopted in a few universities, centers, and labs; it is hoped that similar guidelines will become more widespread.

Applied Research. There are two types of applied research: theory into practice (TIP) and practice into theory (PIT). TIP takes its name from the fact that it begins with theory and asks: How can this be applied to practice? TIP researches a school relevant task such as summarizing or analogies, develops an instructional treatment, and then assesses the effects, using basic research methods, sometimes in the laboratory and sometimes in the schools. TIP has three major distinctive features: First, it is driven by theory in the sense that the primary or ultimate goal is to extend theory. Second, because of this, decisions about the exact dimensions of the experimental task, textual materials, test format, and procedures are guided by methodological concerns rather than practice. Third, TIP studies are therefore typically small-scale and well-controlled in terms of the variables tested and the process of implementation. The research model developed by Brown, Campione, and Day (1980) on summarizing is an example of TIP. Another example is the strand of research which seeks to apply the keyword elaboration mnemonic to a broad range of school tasks from learning foreign languages (e.g., Raugh & Atkinson, 1975) to learning difficult-to-learn definitions and name-fact pairs (e.g., Jones & Hall, 1979; Levin & Shriberg, 1980).

Like basic research paradigms, TIP research would be more relevant to schools if the tasks were more frequently school tasks rather than contrived paradigms and if subjects more often represented problem learners of school ages. Moreover, effective instruction in an experimental lab or suburban school may be less than effective in urban classrooms due to factors in the teaching/learning situation.

To elaborate: What we are suggesting here is not simply "considering" the classroom situation in the early conceptualization of the instruction. Rather, it would be important to specify exactly the conditions under which the instruction should be implemented, and then simulate or provide those conditions. Thus, conditions such as class size, classroom organization in terms of homogeneous vs heterogeneous groupings, tolerance of deviations from the prescribed instructional

guidelines or script, and the like would be specified and controlled for in the same way that sampling and testing procedures are specified and controlled for. Additionally, measures of implementability would have to be developed and used along with, or after, measures of the initial effects of the instruction. This means that an instructional program would have to attain prespecified levels of implementability before being judged as effective to the highest degree. It would be useful to show, for example, that a diverse group of teachers could implement the instruction with gains and variances that were equal to those attained by the experimenter.

The second applied research paradigm, PIT, has existed for some time in schools and regional labs. More recently, however, as more attention has been given to the need for research to have an impact on schools, PIT has emerged in a few universities and centers. Typically, PIT research begins with a practical problem; then theory and basic and applied research methods are used to address the problem. Studies by T. H. Anderson (1979), Armbruster and T. H. Anderson (1980), Weinstein and her associates (1979), Sticht (1979), Jones and Hall (1979), Jones, Amiran, and Katims (1980), Meyer (1980), and Waller and his associates (e.g., Waller, 1980) are examples of this type of research. In each case, the researcher began with a practical question regarding instruction (e.g., what would be the most effective instruction for multiple choice tests?); developed and tested a long-term instructional program in closely-observed treatment/control field implementations in schools, military bases, or universities; assessed the instruction by traditional statistics; and then related the results to both theory and practice. Thus, in contrast to TIP research, PIT research is driven by practice in that its primary goal is to develop instruction that may be used in schools. Consequently, decisions about the task, procedures, and materials are guided by practical considerations rather than those of theory or basic research criteria. PIT studies therefore are typically large-scale,

involving a multiplicity of theories and test variables as well as a lack of control over the way the teachers implement the instructional program. The similarities and differences between TIP and PIT are illustrated in Table 3.

While PIT research is obviously oriented to benefit schools by definition, further gains could be attained if PIT research were extended to address systematically the problems of evaluation referred to above. More specifically, we are referring to developing and testing new measures to evaluate initial program effects and implementability (using new implementation groups) for large-scale programs that are actually used in the schools. Also, we think that paired lab/school research is an interesting solution to the problem of evaluating large-scale programs. T. H. Anderson (1979) and Armbruster and T. H. Anderson (1980), for example, developed and field tested an innovative form of outlining to be used in schools. In addition to gathering data in the schools, they tested the method using sample segments of text in a well-controlled study.

A related problem concerns the effects of large-scale programs (or their component parts) on cognition and cognitive development. Currently, reading instruction in the schools is based on commercially produced basal systems which have no theoretical basis (e.g., Harcourt's Bookmark Program, 1979) commercially produced systems which purport to be based on cognitive theory (e.g., Open Courts Cognitive Series, 1978), programs developed by schools which purport to address problems of cognition and cognitive development (e.g., CMLR), and programs developed by schools which do not in any way attempt to address problems of cognition and development (e.g., any number of mastery learning programs which are developed by teachers and curriculum staff who have very limited research skills and training in cognitive theory). Most of the commercially developed programs have no data at all, not even field test data, to document their effects on comprehension and learning. And while most of the school-based programs have field test data and/or program evaluation data, none of them

to the knowledge of the authors have any data on the effects of the program as a whole, or parts of it, on any of the variables of concern to cognitive psychologists. There is no information, for example, on the effects of these programs on measures of comprehension other than standardized tests and criterion-referenced tests which typically have not been validated by a research procedure other than judgment by a panel of teachers.

PIT research could also be extended by more effort to use practice to refine theory. A case in point is described by Jones, Amiran, and Katims (1980). The discovery that they made in the field--that the most effective instruction involved specifying to the student the structural features of the text--caused them to re-examine the existing research literature. This led to a second discovery. Even for such simple tasks as categorizing or elaboration, ineffective instruction was often associated with failure to specify the structural features of the text (e.g., Schultz & Di Vesta, 1972). Obviously, effective instruction (for low achieving students or younger students) usually involved explicit instruction in the structural characteristics as well as the process of applying the strategy - in spite of the fact that all of the studies examined purported to be testing a single variable (learning strategy instructions). This re-examination eventually led to the development of taxonomies of learning strategies task attributes and text structures used in the research literature. These taxonomies may be used to control the features of learning strategies instruction and text structures that appear in any given instructional text. They may also be used for basic and applied researchers to specify what dimensions of a strategy, task, and test are controlled for in a given experiment.

Political Issues

If the proposed changes are to take place, there would have to be major changes in the system of rewards for funding research and development projects, the awarding of positions in schools and universities, and available forums to

disseminate knowledge. First, it would be important for funding agencies (1) to fund more research that is relevant to instruction in schools; (2) to require that qualified school staff act as consultants, liaison staff, or collaborative researchers to define what is school relevant and/or to develop the instruction; (3) to require the authors of research purported to be school relevant to specify how it is school relevant, to specify the full details of the teaching/learning situation under which it was tested, and to provide evidence of its implementability by teachers other than those in the initial experiment; (4) to fund more university/school collaborations whereby the university provides technical assistance to schools to improve instruction in ways jointly defined by the collaborating parties, including the selection of staff participating in the collaboration; (5) to encourage university staff to gain a working knowledge of how schools operate and the constraints of the teaching/learning situation on instruction; (6) to require schools seeking funding for instructional development to hire staff, with a research background in reading, learning, instruction, curriculum design, cognitive psychology, and/or educational psychology, as well as trained research staff; (7) to fund research involving laboratory tests of research-based instruction developed in the schools; (8) to devote considerably more research funds to schools or school/university collaborations to develop cognitive-based instruction and to assess the effects of both cognitive-based and non-cognitive based instruction on cognition, learning, and development; and (9) to require researchers to disseminate their results in forums that are accessible to school administrators and teachers using language that school personnel can understand.

Universities and centers could (1) encourage individual professors to initiate school relevant research, including laboratory tests of existing instructional units and the evaluation of large scale programs, (2) encourage methodologists to develop and test measures of evaluation which are more appropriate for use in large schools, (3) establish more graduate courses and/or doctoral programs in the

area of instruction, and curriculum design within research-oriented departments, (4) hire qualified school staff to act as consultants to research projects and course planning as well as team partners for courses on school/classroom constraints on learning, cognition, development, and instruction. In turn, schools could (1) hire more university staff to instruct curriculum developers, school researchers, and administrators about development in research or instruction, curriculum design, learning, reading, cognition, and development; (2) hire more staff with research backgrounds to develop and evaluate curriculum and instruction; and (3) encourage appropriate school staff to inform themselves about such matters by greater participation in professional organizations or divisions devoted to research or instruction and related topics. Finally, professional organizations which sponsor research journals would need to include a class of articles which are currently excluded. This refers to the inclusion of school-based articles to journals whose subscribers are primarily university researchers as well as an effort on the part of university and school researchers to communicate to administrators, curriculum developers, and teachers about research relating to learning and instruction. This suggestion applies to descriptions of research-based programs, evaluations of such programs, and integrative articles relating research to schools and school problems to researchers. Generally speaking, these additions should be inclusions of well-established journals, rather than diverted to new journals, so that they will be read by the appropriate audiences.

Most of these suggestions are not new. Almost all of them are based on exemplary policies recently instituted at the National Institute of Education and a small number of academic institutions and schools. They are given here not to be original but rather to reinforce and extend. At the moment, universities and schools are only at the very beginning stages of courtship; it is hoped that the above suggestions greatly facilitate the process of both partners getting to know each other so that ED can flourish as an increasingly effective member of the research community.

Notes

1. At first glance problems 5 and 6 appear to be contradictory but they are not; they refer to different stages of implementation. Initial implementations are likely to involve large variances from individual differences in student performance and differences in teaching, but these decrease greatly with continued use of the program.

References

Anderson, T.H. Study skills and learning strategies. In H.F. O'Neil & C.D. Spielberger (Eds.), Cognitive and attentive learning strategies. New York: Academic Press, 1979.

Andre, M.E.D.A., & Anderson, T.H. The development and evaluation of a self-questioning study technique. Reading Research Quarterly, 1979-79, 14, 605-623.

Armbruster, B.B., & Anderson, T.H. The effect of massing on the free recall of expository text (Tech. Rep. No. 160). Urbana: University of Illinois, Center for the Study of Reading, February 1980.

Bereiter, C., Hughes, A., & Anderson, V. Catching On: Levels A-E (Open Court Cognitive Series: Reading Comprehension). LaSalle, Ill.: Open Court, 1977.

Benjamin, R. Making schools work. New York: Criterion Press, 1981.

Brown, A.L., Campione, J.C., & Day, J.D. Learning to learn: On training students to learn from texts. Educational Researcher, December 1980.

Brown, A.L., & Day, J.D. Strategies and knowledge for summarizing texts: The development of expertise. Manuscript submitted for publication 1980.

Carnine, R. & Silbert, J. Direct instruction Reading. Charles E. Merrill Publishing Co., 1979.

Chicago Board of Education, Chicago Mastery Learning Reading. Watertown, Mass.: Mastery Education Corp., 1980.

Durkin, D. What classroom observations reveal about reading comprehension instruction. Reading Research Quarterly, 1979-79, 14, 481-533.

Englemann, Z. (Ed.) DISTAR. Chicago: Science Research Associates, 1978.

Hansen, J. & Pearson, P.D. An instructional study: Improving the inferential comprehension of good and poor fourth-grade readers. In press.

Harcourt, Brace, & Jovanovich, Bookmark Reading Program. New York: Harcourt, Brace & Jovanovich, 1979.

Jones, B.F., Amiran, M.R., & Katims, M. Embedding structural information and strategy instructions in reading and writing instructional texts: Two models of development. Paper presented at the National Institute of Education and Learning Research and Development Center, University of Pittsburgh, conference on Thinking and Learning Skills, October 1980.

Jones, B.F., & Hall, J.W. School applications of the Key word method as a study strategy of eighth graders. Paper presented at the annual meeting of the American Psychological Association, Toronto, August 1979.

Jones, B.F., & Hall, J.W. Effects of cross-classification strategies for recalling prose and writing compare-and-contrast essays. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 1979. ED 182 712.

Jones, B.F., & Katims, M. Chicago's mandatory summer schools: The bottom-line of a strict, mastery-based promotion policy. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Jones, B.F., Monsaas, J.A., & Katims, M. Improving reading comprehension: Embedding diverse learning strategies within a mastery learning instructional format. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 1979.

Joyce, B., & Showers, B. Inservice programs: The messages of research. Educational Leadership, 1979.

Katims, M. Implementing mastery learning reading instructions in the Chicago Public Schools. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 1979.

Katims, M., & Jones, B.F. Class effects and school effects: Reading improvement in an inner city school. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, April 1981a.

Katims, M., & Jones, B.F. Chicago Mastery Learning Reading: A Technology for School Improvement. Educational Technology. In press.

Katims, M., Smith, J.K., Steele, C., & Wick, J.W. The Chicago Mastery Learning Reading Program: An interim evaluation. Paper presented at the annual meeting of the American Educational Research Association, New York, April 1978.

Meyer, B.J.F. Signaling in Text. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Rand Corporation, Change Agent Study (Vol. 4). Santa Monica, Calif.: Rand Corp., 1978.

Raugh, M.R., & Atkinson, R.C. A mnemonic for learning a second language vocabulary. Journal of Educational Psychology, 1975, 67, 1-16.

Reid, E. Teaching Reading Comprehension. Salt Lake City: Cone, 1978.

Rohwer, W.D., Jr. Prime time for education: Childhood or early adolescence. Harvard Educational Review, 1971, 41, 316-341.

Rosenshine, B.V. Content, time, and direct instruction. In P.Peterson & H. Wahlberg (Eds.), Research on teaching: Concepts, findings, and implications. Berkeley, Calif.: McCutcheon, 1979.

Schultz, C.R., & DiVesta, F.J. Effects of passage organization and note taking on the selection of clustering strategies and on recall of textual materials. Journal of Educational Psychology, 1972, 63 244-252.

Shriberg, L.K., & Levin, J.L. Learning about famous people via the key word method. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Sternberg, R.J. Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities. Hillsdale, N.J.: Erlbaum, 1977.

Sticht, T. Developing literacy and learning strategies in organizational settings. In H.F. O'Neil, Jr., & C.D. Spielberger (Eds.), Cognitive and affective learning strategies. New York: Academic Press, 1979.

Stoll, L.J. A model for sequencing skills instruction. Paper presented at the annual meeting of the International Reading Association, St. Louis, May 1980.

Waller, R.H. Typography and design of text. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 1980.

Weinstein, C. Elaboration skills as a learning strategy. In H.F. O'Neil (Ed.), Learning strategies. New York: Academic Press, 1978.

Table 1

Comparison of Successful and Unsuccessful Teachers

	Reading Comprehension	Mathematics Problem Solving
Successful Mastery Learning Reading Instruction	M = 1.39 SD = .38	M = .84 SD = .44
Unsuccessful Reading Instruction	M = .36 SD = .20	M = .73 SD = .26

Table 2

Means and Standard Deviations of Age Cycle Gains

Year	Mean	Standard Deviation
1977	.80	.558
1978	.57	.578
1979	.73	.440
1980	1.03	.287